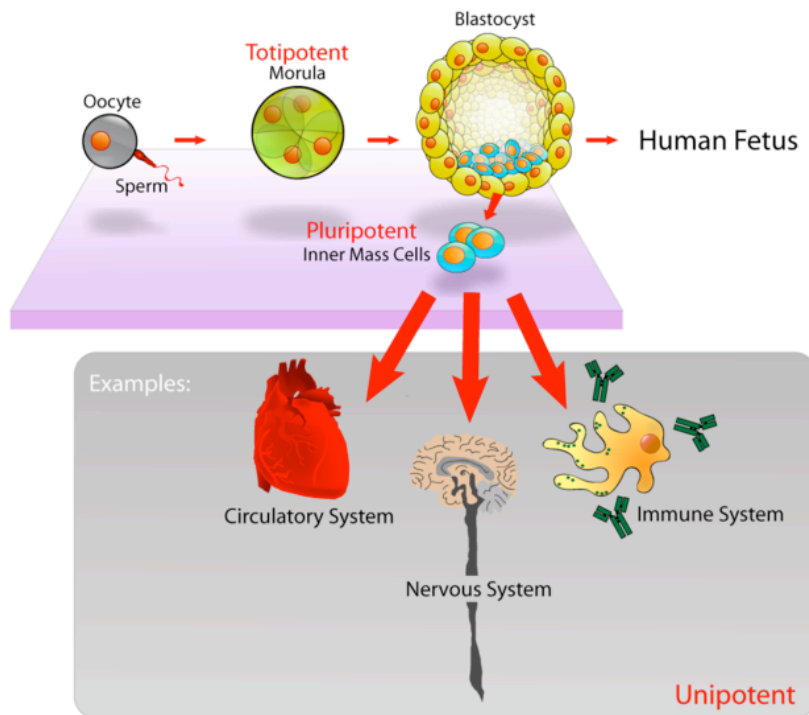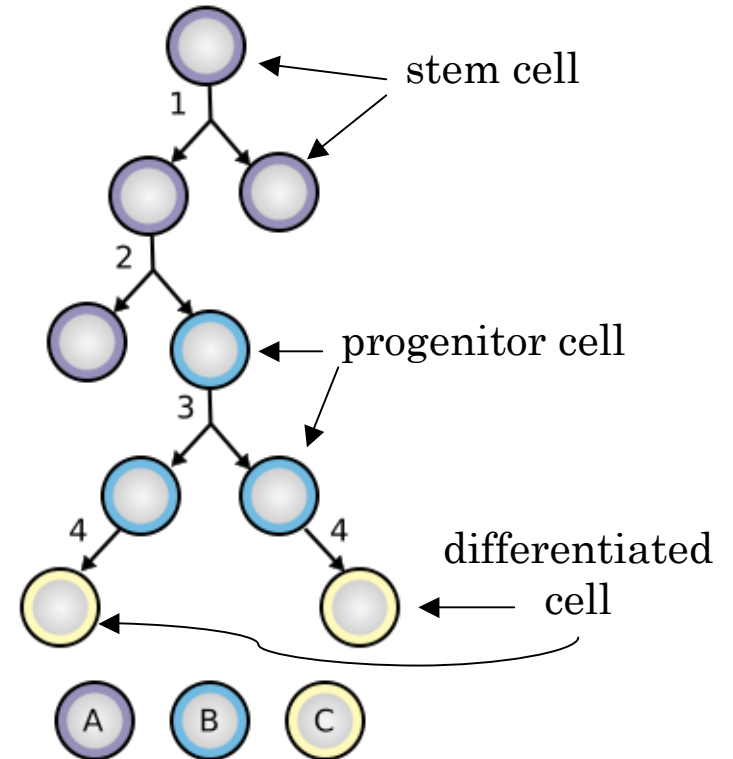# Computational learning of stem cell fates

Martina Koeva

09/10/07

The fascinating world of stem cells

- Adult and embryonic stem cells
- Pluripotency and multipotency
- Differentiation and proliferation

http://en.wikipedia.org/wiki/Image:Stem_cells_diagram.png

http://en.wikipedia.org/wiki/Image:Stem_cell_division_and_differentiation.svg

Therapeutic potential of stem cells
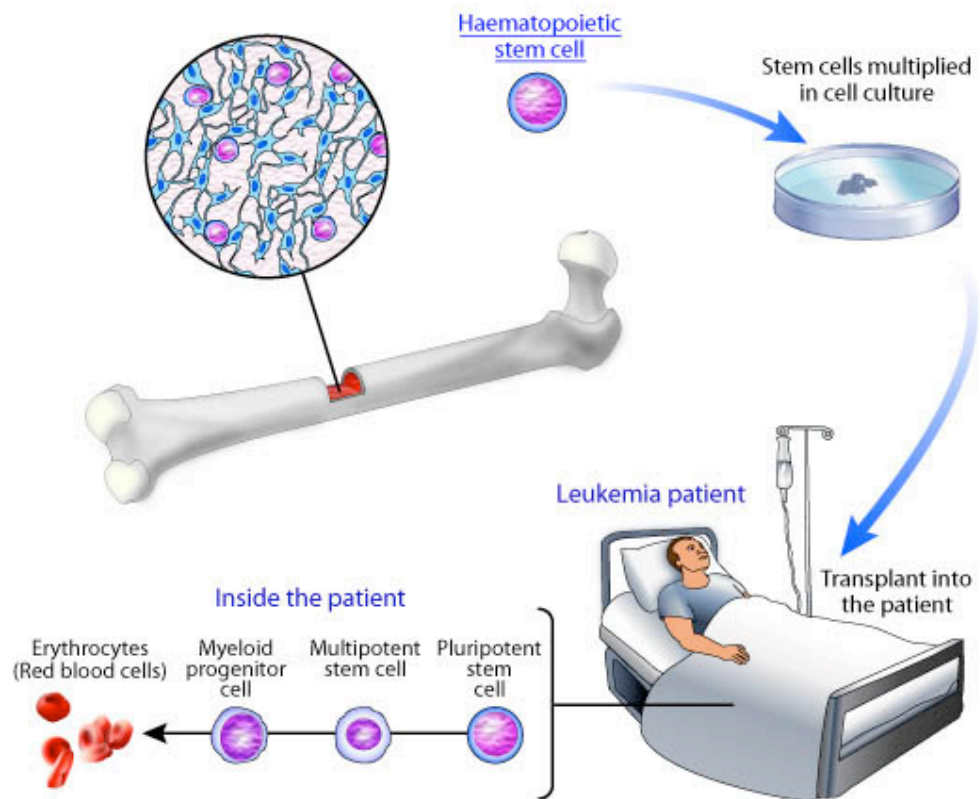
- Parkinson's disease
- Cancer
  - leukemia

Illustration by Cell Imaging Core of the Center for Reproductive Sciences.

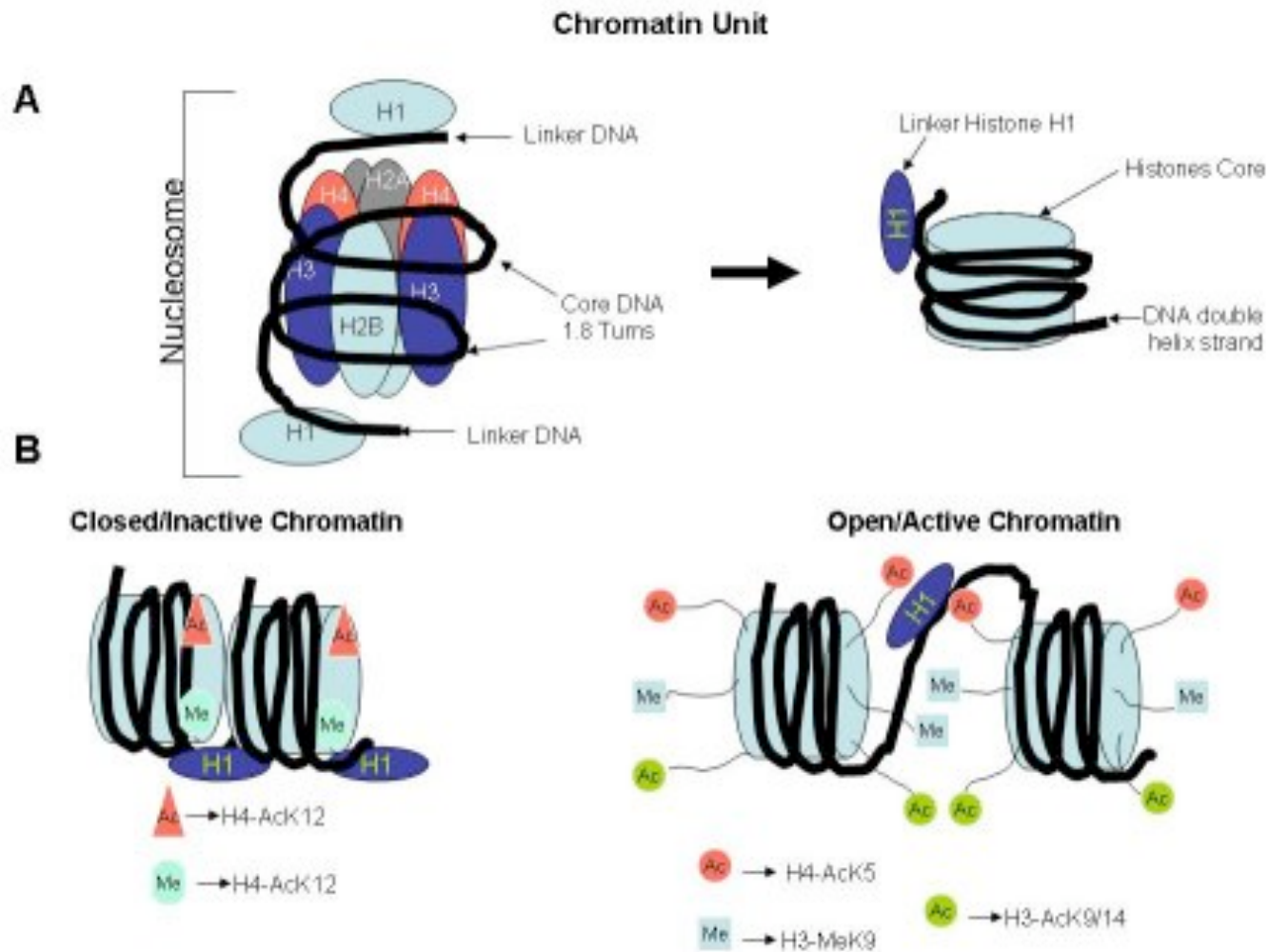http://www.kumc.edu/stemcell/mature.html

Current challenges in stem cells

- Chromatin, chromatin state and differentiation
- MiRNAs and differentiation
- More and better marker genes

Proposed aims

- Aim 1: Assess coherence of gene modules in stem cell differentiation
  - Chromosomal gene neighborhoods
  - Predicted targets of a miRNA

- Aim 2: Identify and classify cell state in stem cell differentiation using gene expression data

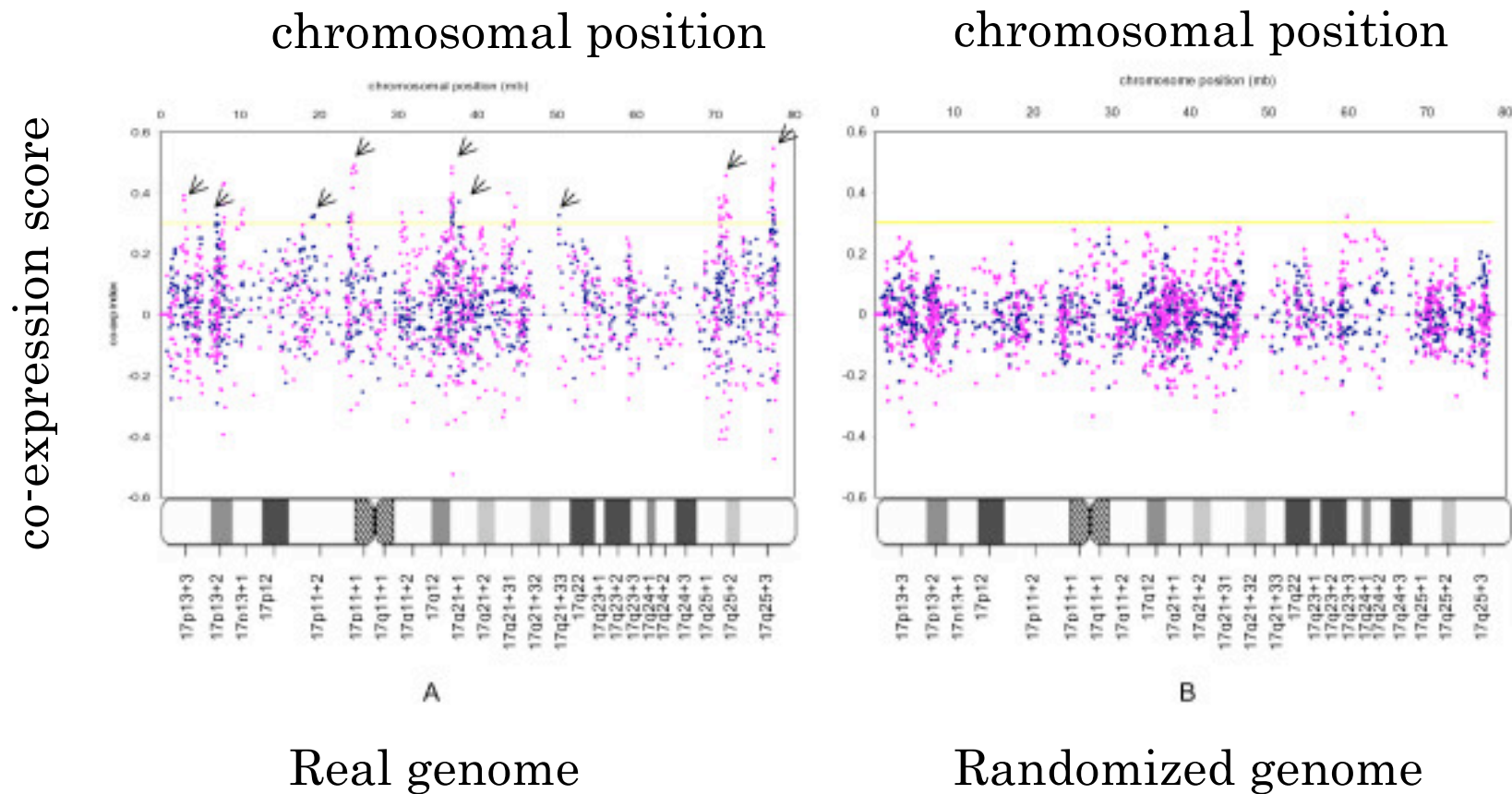- Aim 3: Identify differential gene expression patterns in hierarchical stem cell lineages
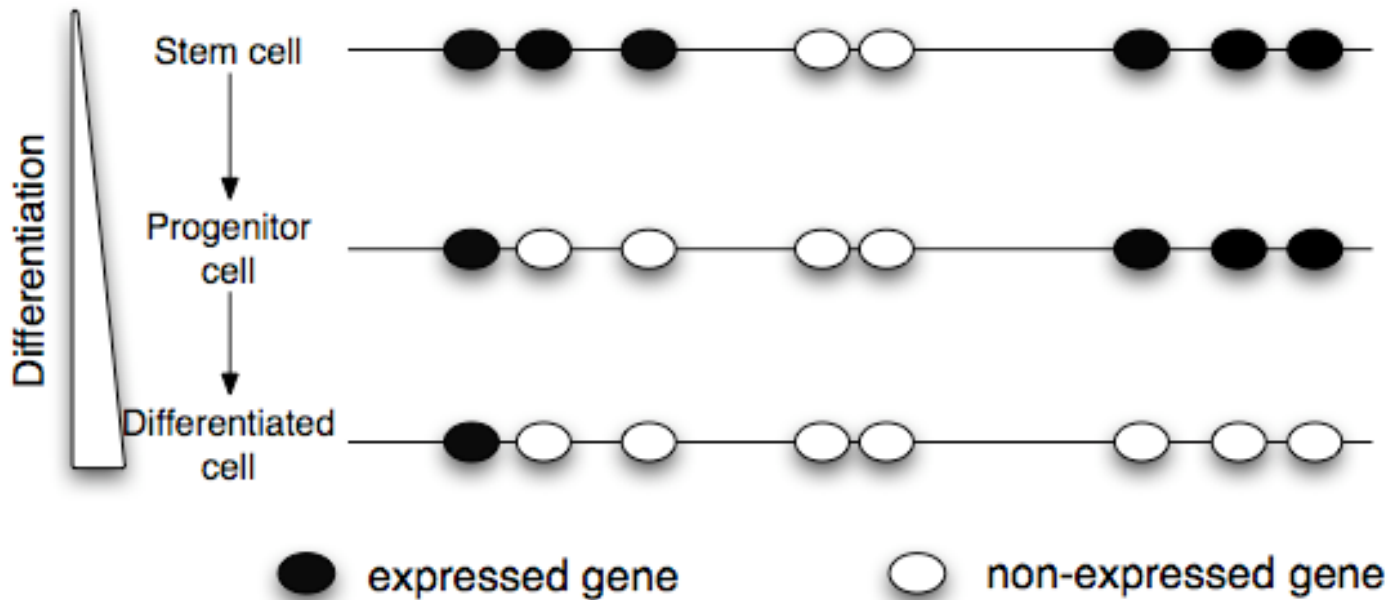
Open and closed chromatin

Stem cells show domains of co-expression on the chromosome

Fig.2

chromosomal position

chromosomal position

co-expression score

A

B
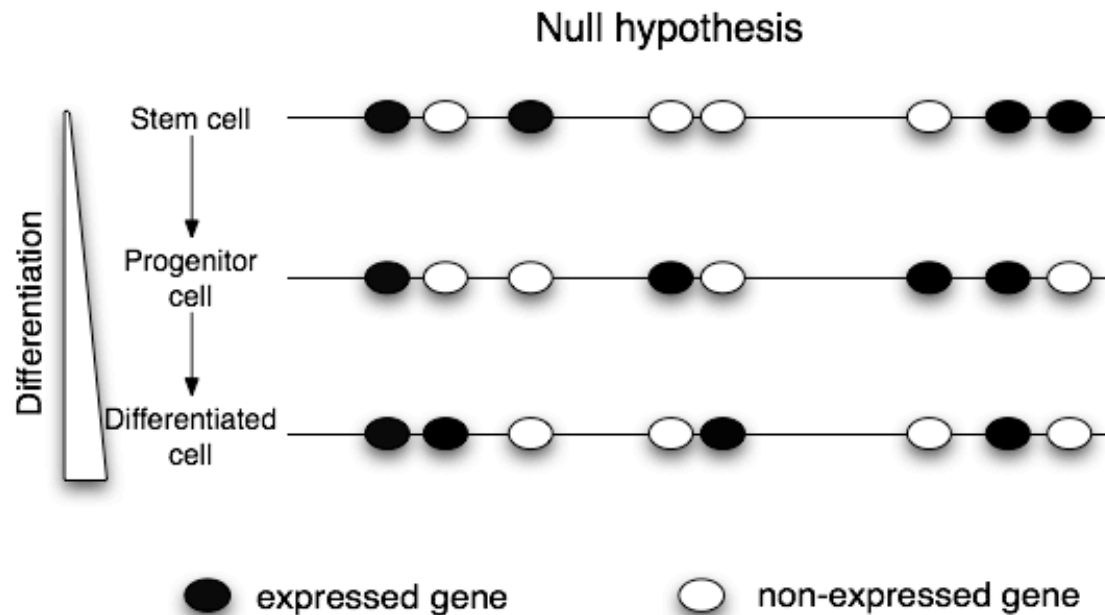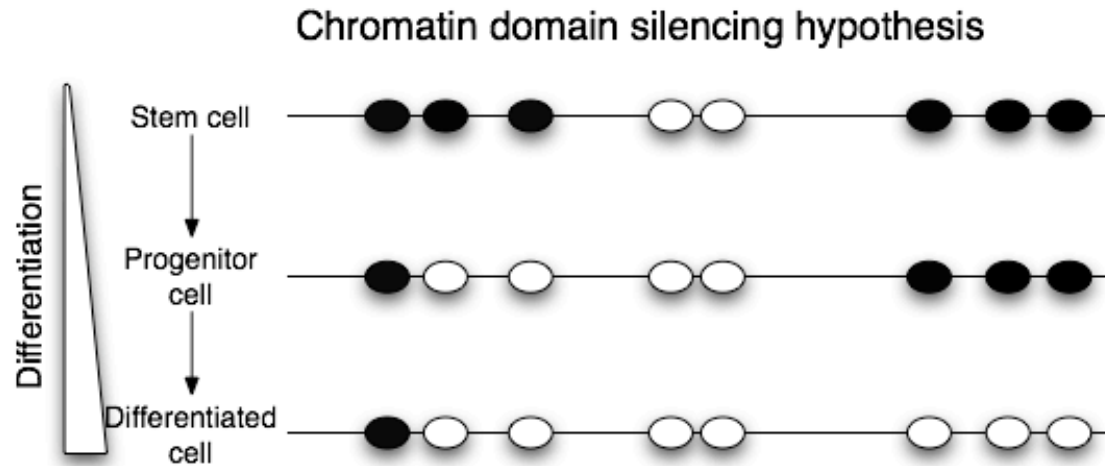
Real genome

Randomized genome

Li 2006

Aim 1: Test domain silencing hypothesis

- Stem cells - "open" chromatin
- Differentiation - "closed" chromatin

Chromatin silencing hypothesis

Chromatin domain silencing hypothesis

Null hypothesis

Hematopoietic system in mouse

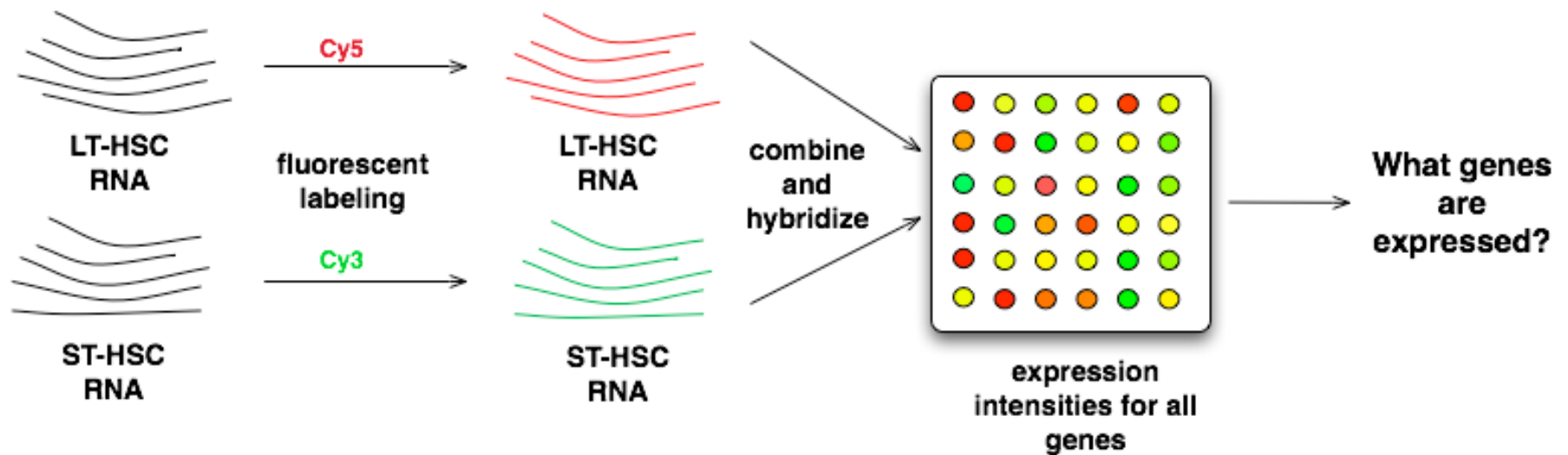http://www.molmed.lu.se/HSC_regulation.htm
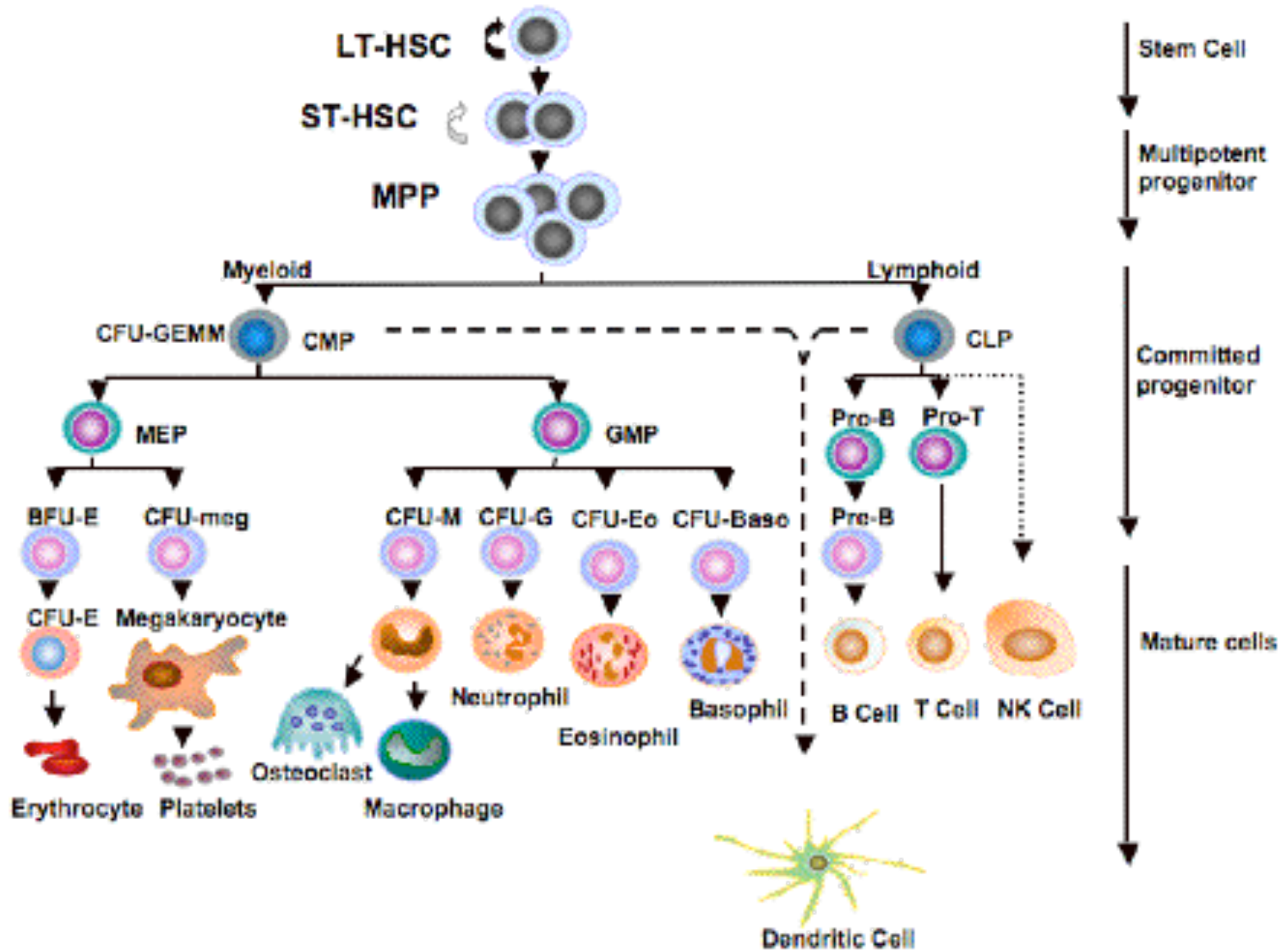
High-throughput gene expression data in the hematopoietic system

- Weissman lab
- cDNA microarray data in mouse
- Pairwise comparisons between LT-HSC, ST-HSC and MPP cell populations
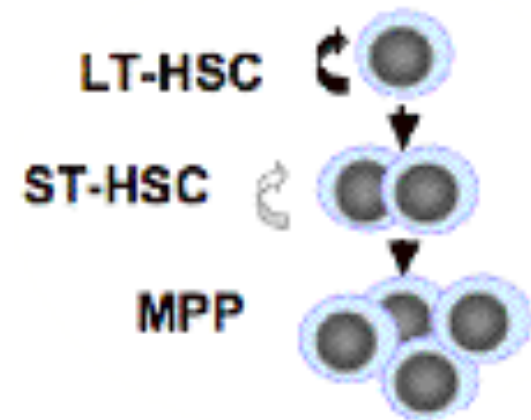
What genes are expressed?

Adapted from
http://www.microarrayworld.com/

- Relative expression between conditions
- Probability of expression of gene in each condition

Empirical probabilistic expression detection

- Probabilistic empirical Bayesian method for expression estimation of a gene
- Positive and negative control distributions
- Average posterior probability for each gene
- Evaluated against an ANOVA FDR-based approach

Global windowing approach

- Probability of co-expression within window
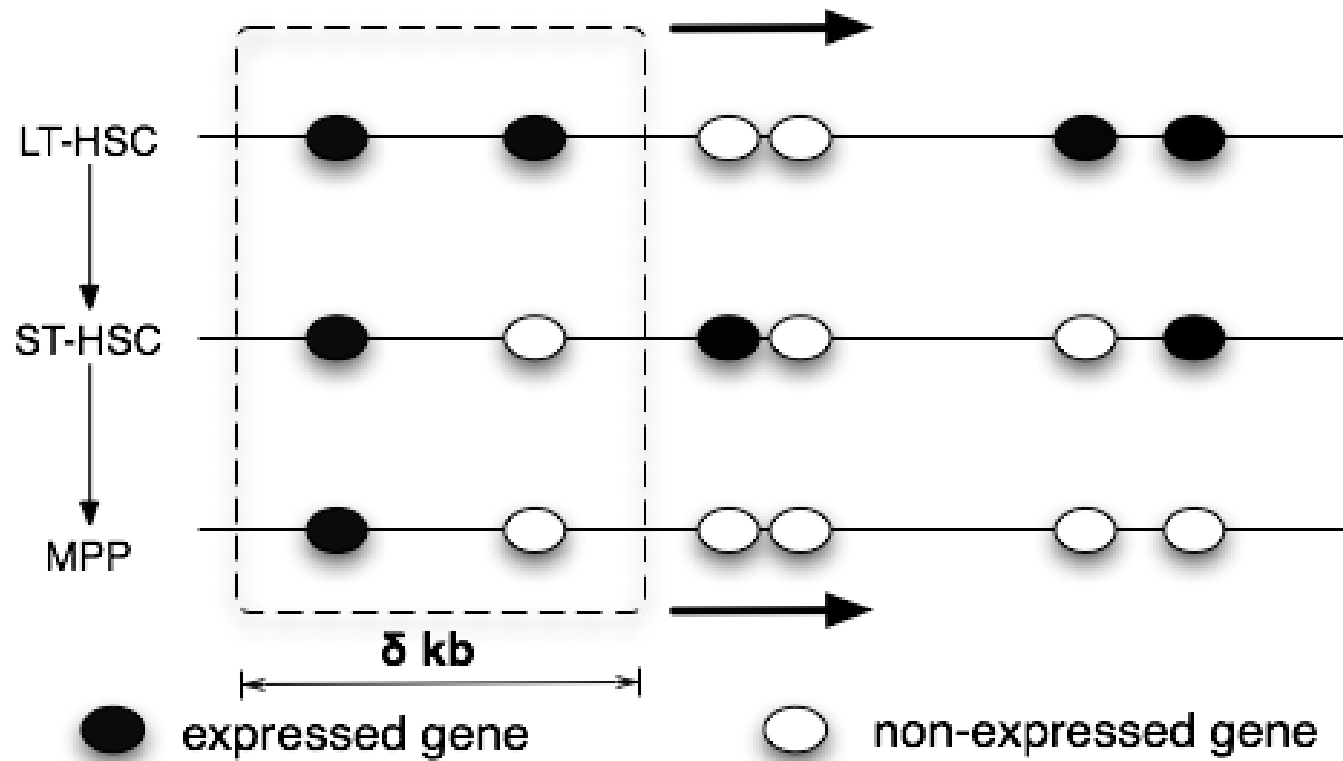
Co-expressed genes within window

$$P\left(g_i = 1 \middle| g_{i+1} = 1, d_{i,i+1} = \delta\right) = \frac{P(g_i = 1, g_{i+1} = 1, d_{i,i+1} = \delta)}{P(g_i = 1, g_{i+1} = 1) \ P(d_{i,i+1} = \delta)}$$

Co-expression of neighboring genes   Genes within distance
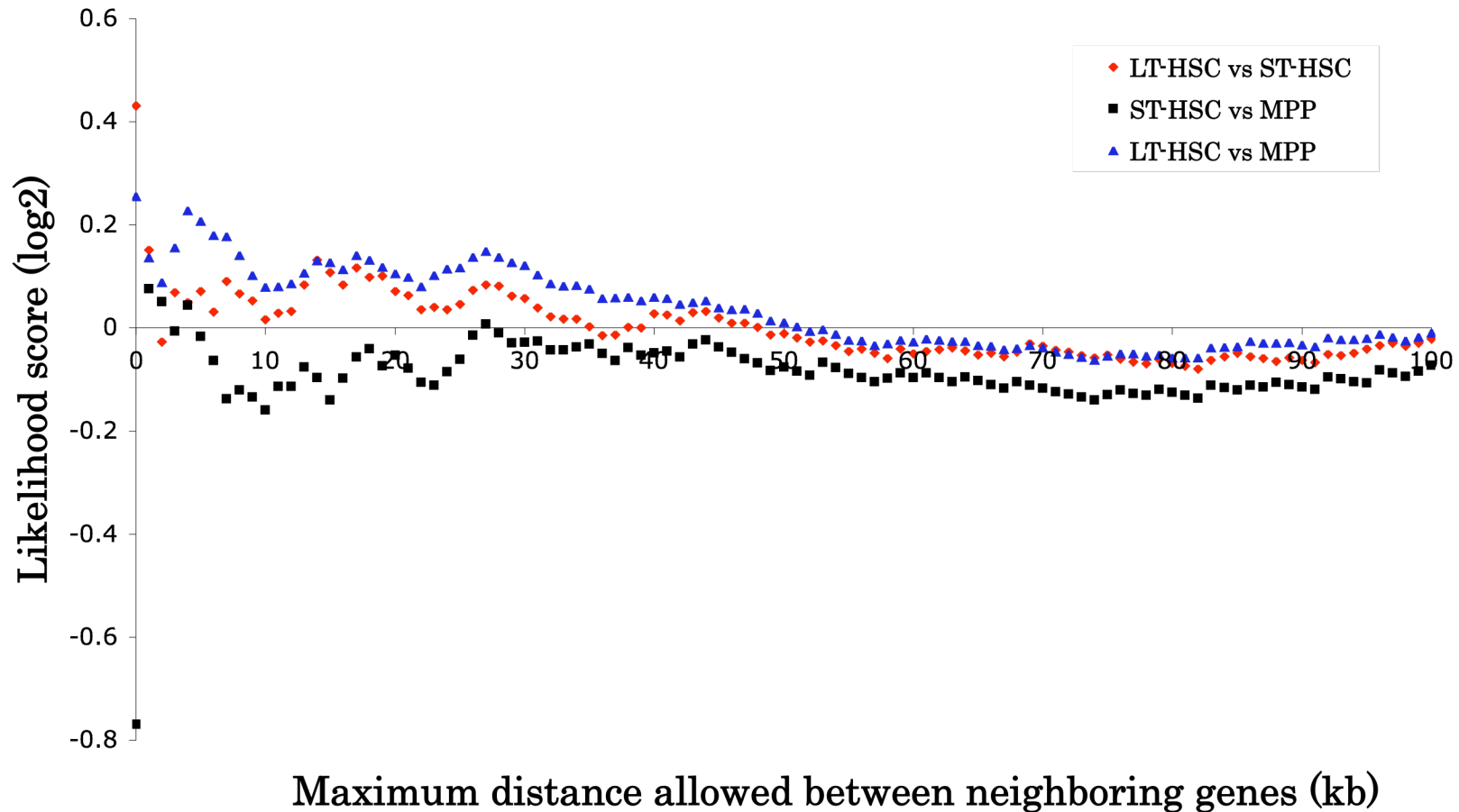
- Global effects
  – Windowing approach - two gene window
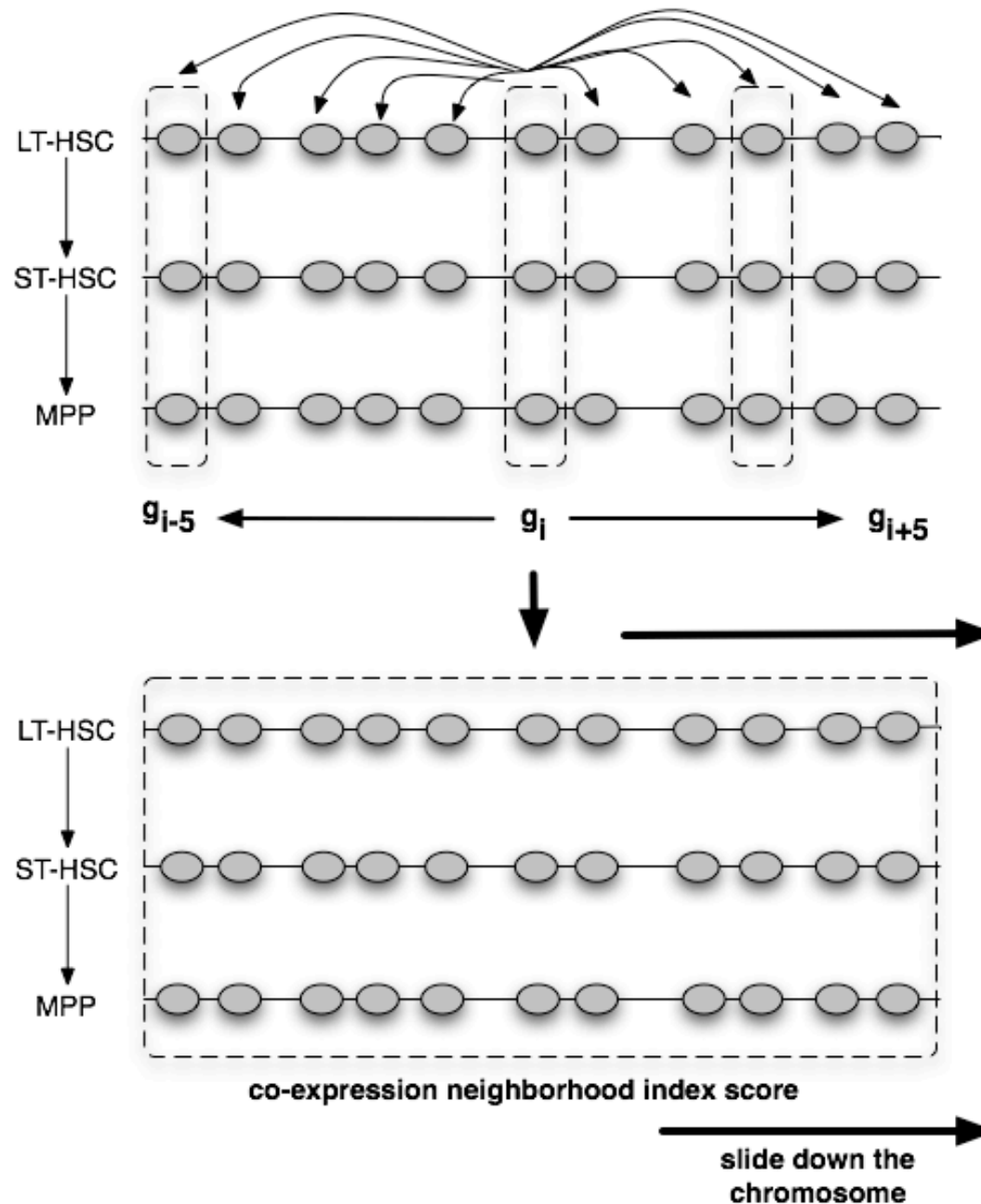 – Likelihood score

Global windowing approach

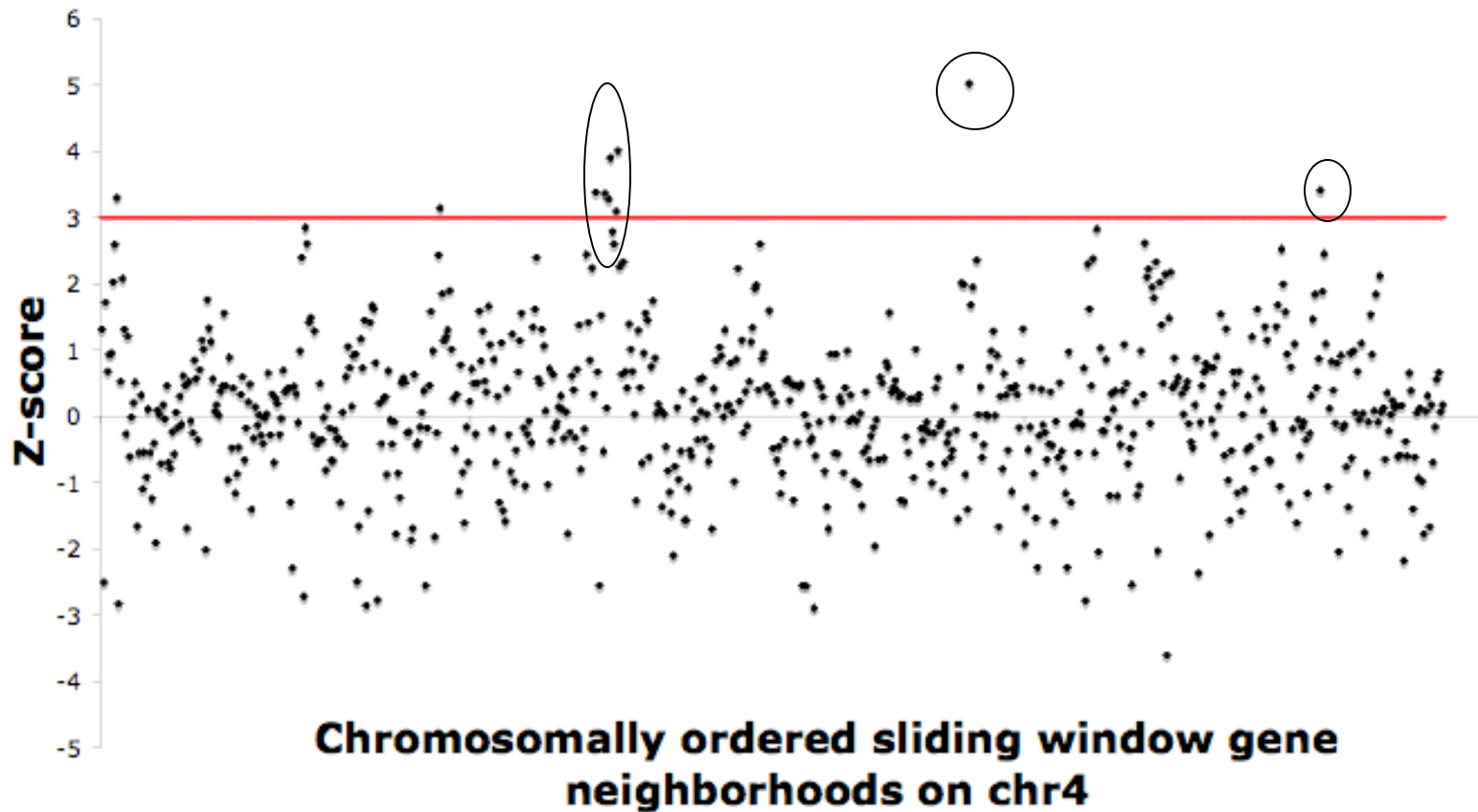Global assessment of likelihood of co-expression of neighboring genes at different distance cutoffs
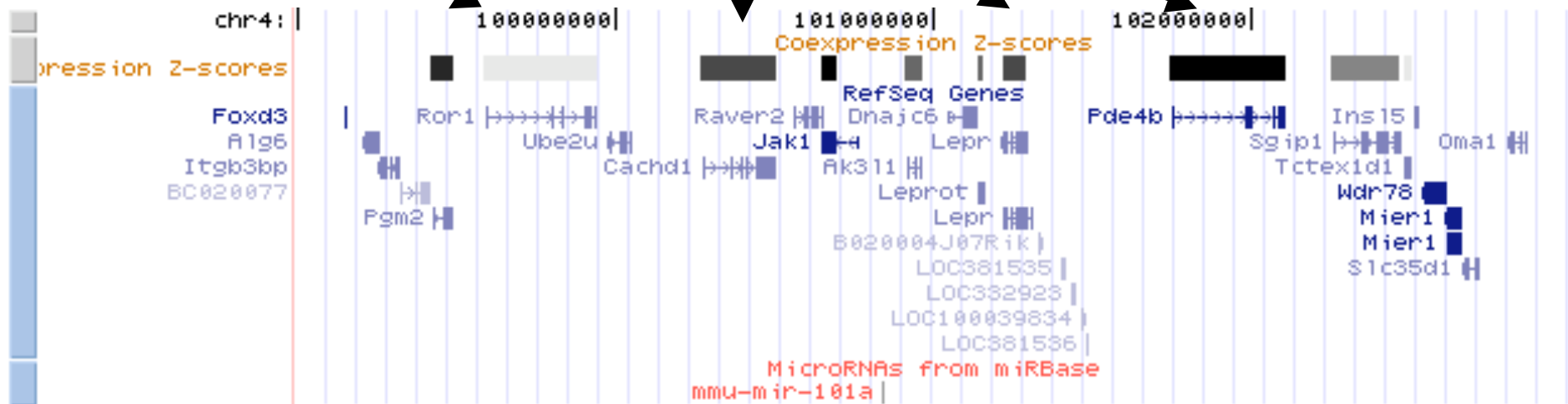
Local windowing approach

Evaluation of co-expression cluster index scores for all gene neighborhoods on chr4 in mouse ES/EB differentiation experiment

Z-score

Chromosomally ordered sliding window gene neighborhoods on chr4

Gene neighborhood with significant co-expression scores


Gene neighborhoods with significant scores

Ror1 - receptor tyrosine kinase
Jak1 - Jak tyrosin protein kinase
Lepr - Leptin receptor precursor
Pde4b, Pgm2

Summary and proposed steps for chromatin domain analysis

- Co-expressed chromosomal gene neighborhoods
  - Identification and evaluation
- Chromatin domain silencing hypothesis
  - Evaluation
- Publicly available stem cell differentiation experiments
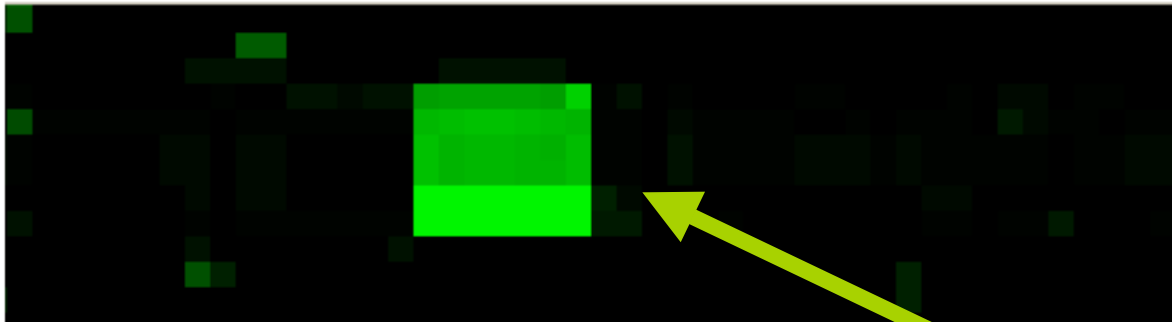
Role of microRNAs in gene regulation
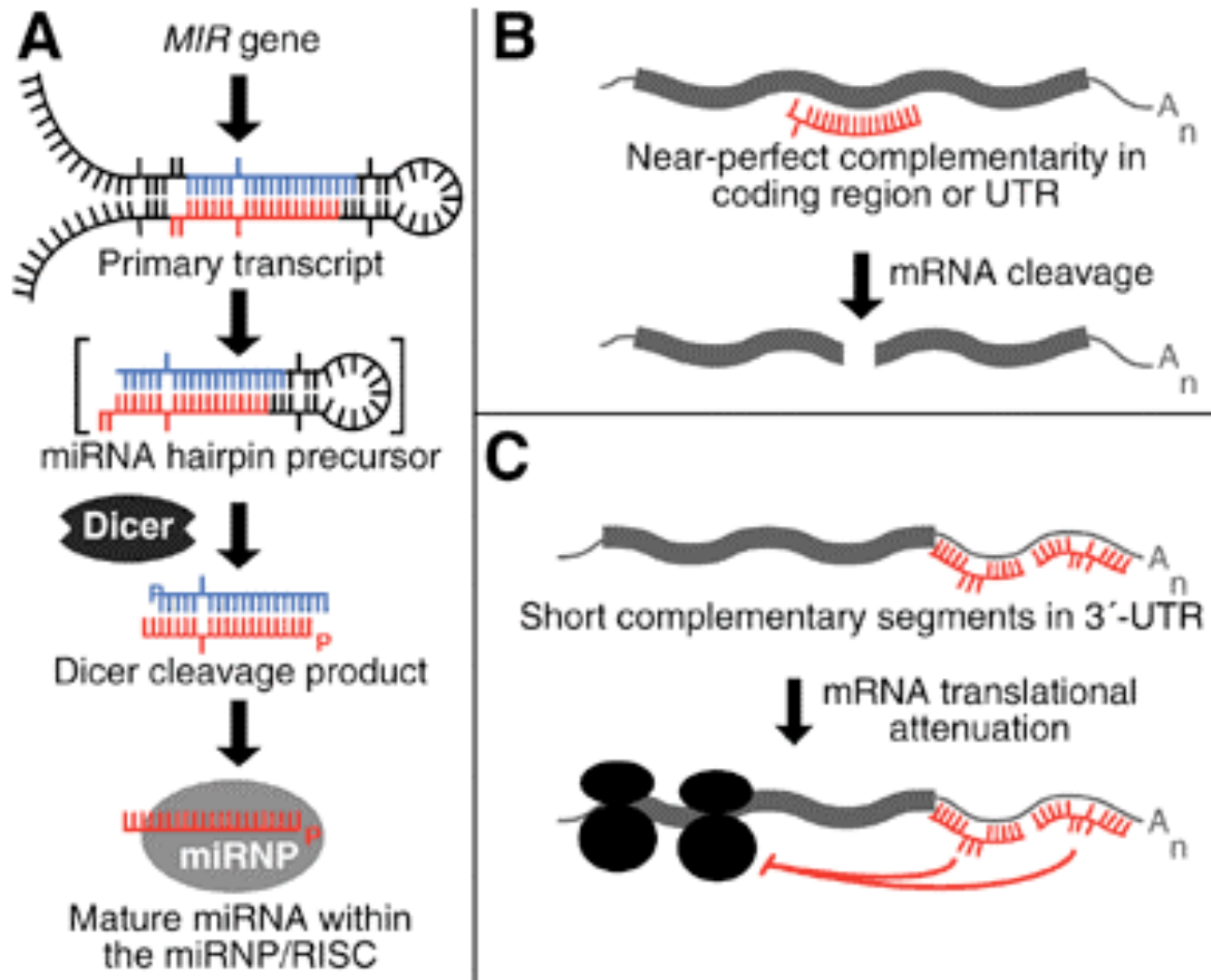
MicroRNAs in the hematopoietic system

- Weissman lab
- Differentially expressed miRNAs in human
  - Hematopoietic system
  - What do they do?
- Prediction of miRNA targets
- Can we tie miRNA expression and miRNA target expression?

Functional enrichment of predicted targets

Cluster of miRNAs differentially expressed between HSCs and LSCs
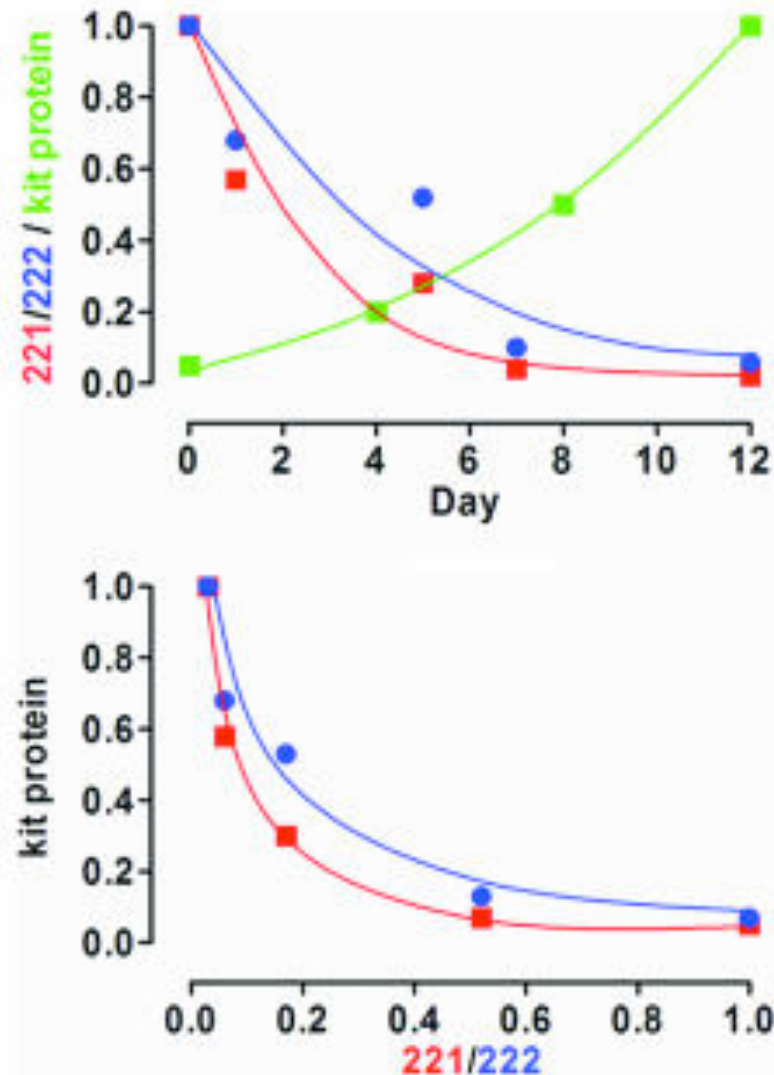
Cell adhesion; Cell-cell adhesion; Calcium ion binding

Daniel Sam

Role of miRNAs in differentiation through target expression analysis

- Predicted targets with similar expression profiles
  - Common regulation

- Conservation of target expression through evolution

MicroRNAs can show inverse correlation to their predicted targets during differentiation
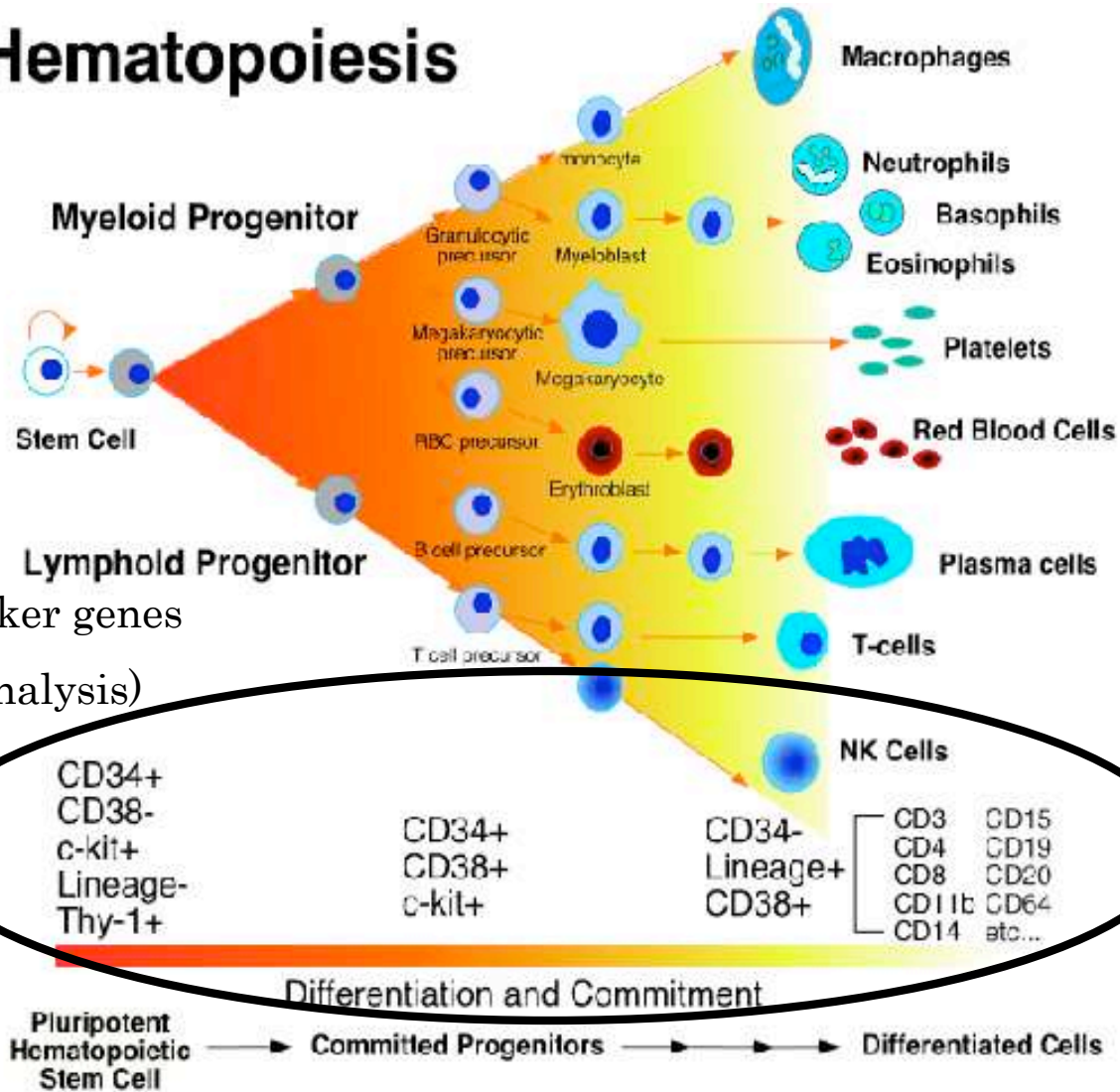

Felli 2005

Summary and proposed steps for miRNA role in differentiation analysis

- Modules of miRNA targets with shared expression profiles
    - Identification and evaluation
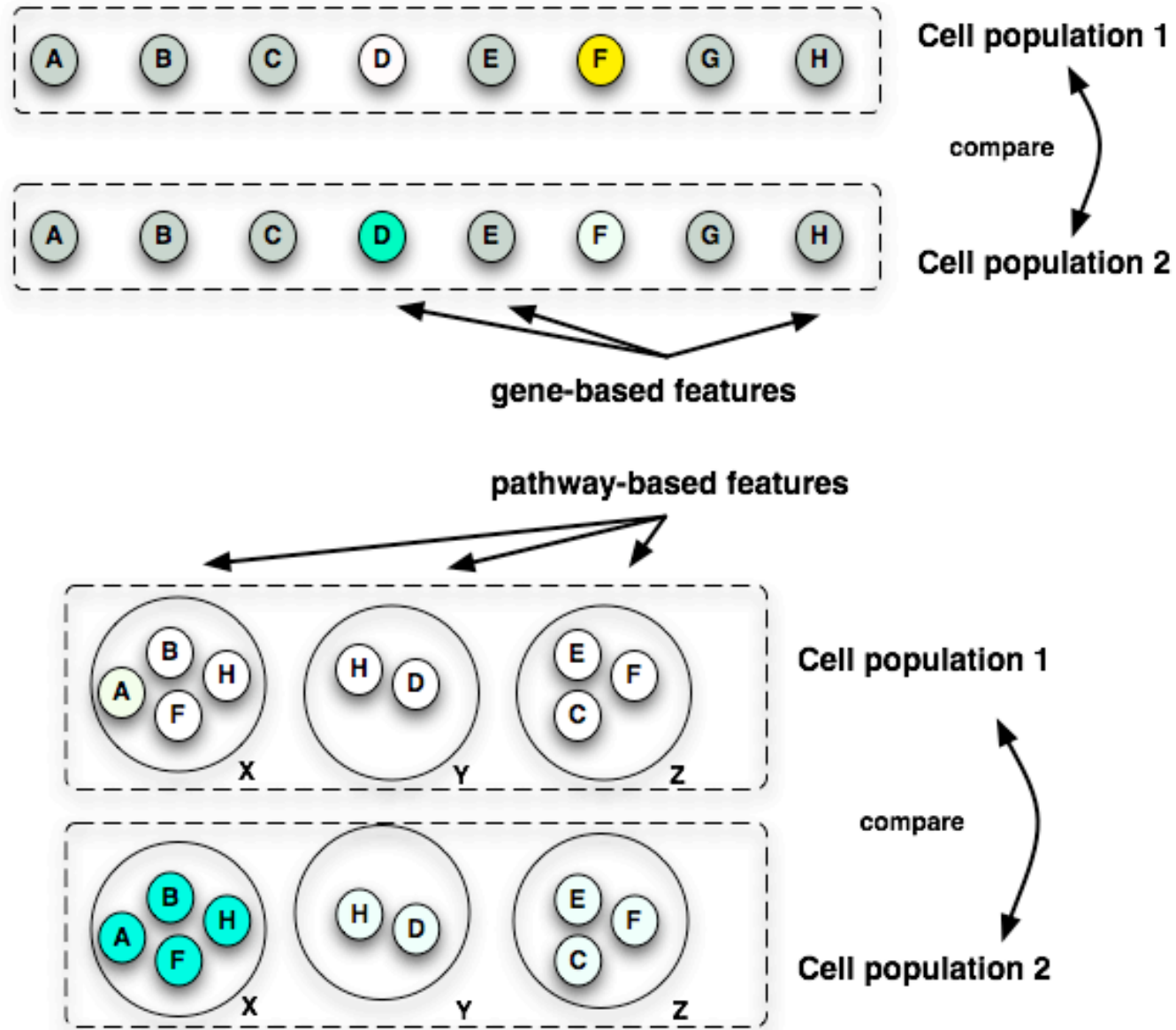- Role of specific miRNAs in differentiation
    - Evaluation

Stem cell state classification

Cell surface marker genes
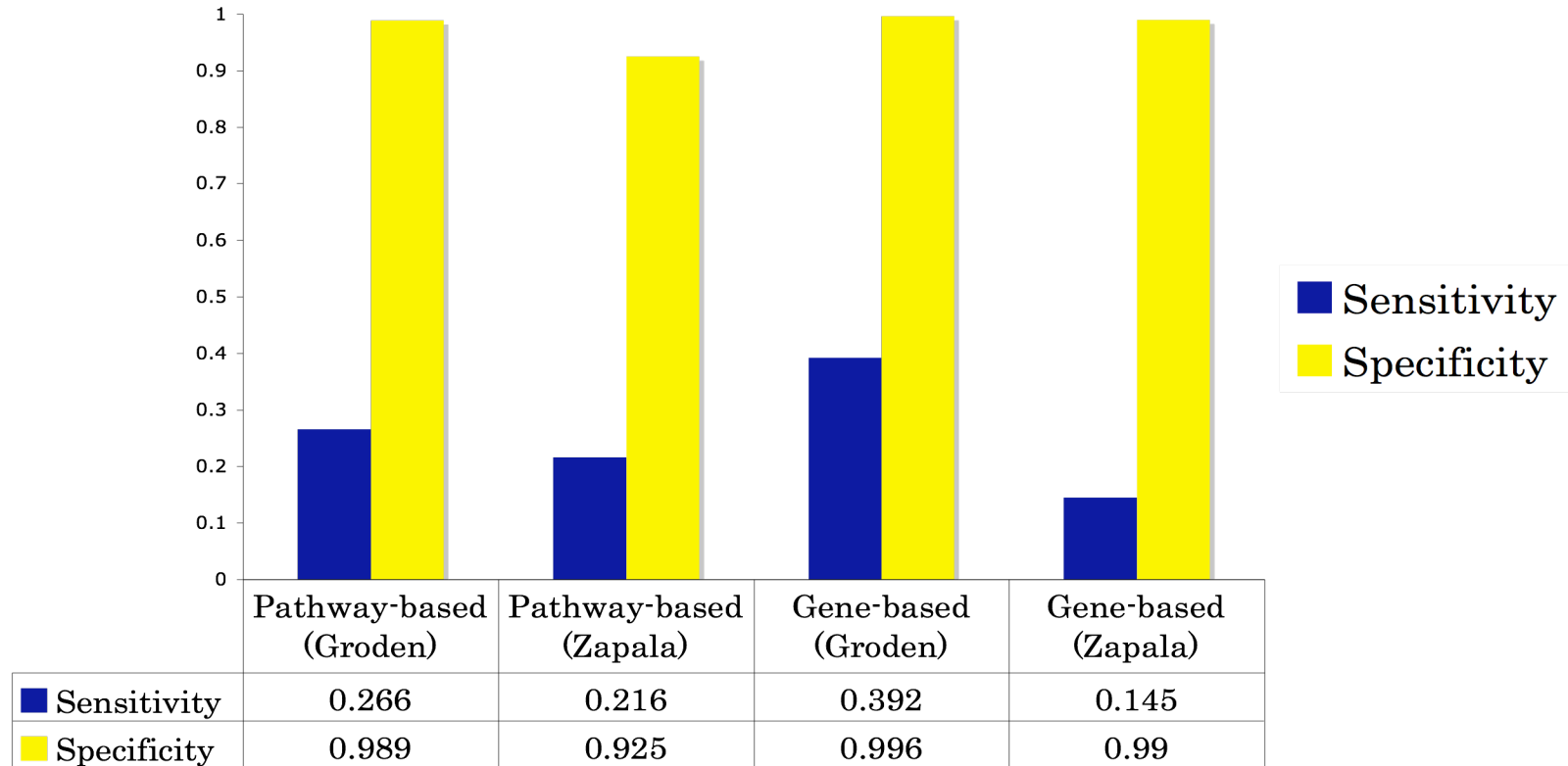(used in FACS analysis)

Gene-based and pathway-based features

Tissue comparisons using different feature types

	Pathway-based (Groden)	Pathway-based (Zapala)	Gene-based (Groden)	Gene-based (Zapala)
Sensitivity	0.266	0.216	0.392	0.145
Specificity	0.989	0.925	0.996	0.99

Aim 2: Classify cell state in differentiation experiments
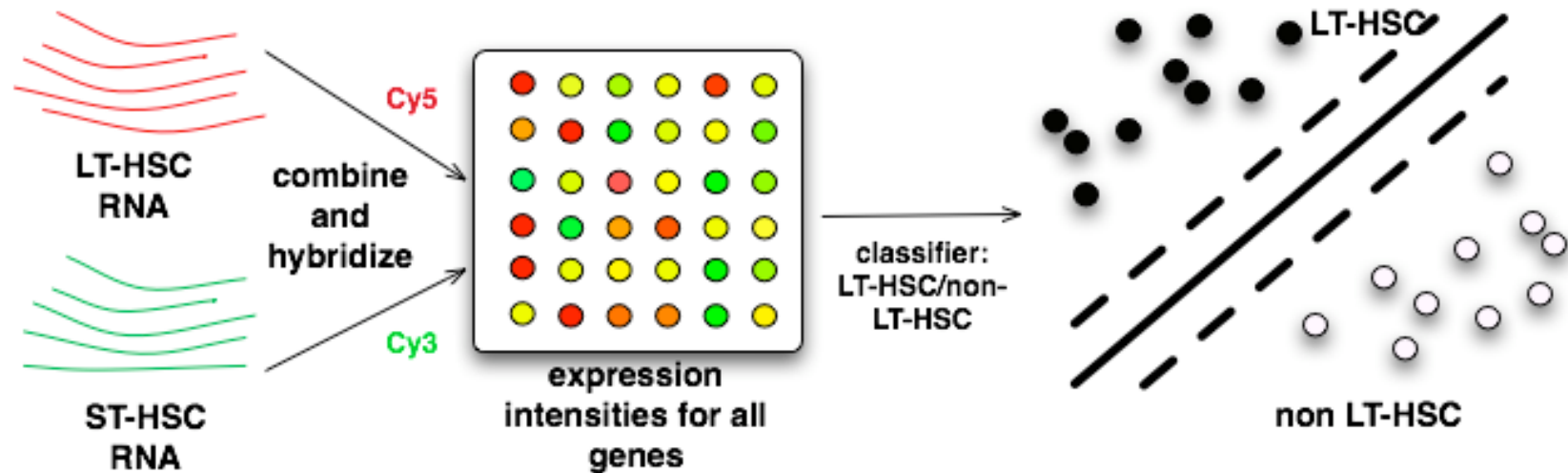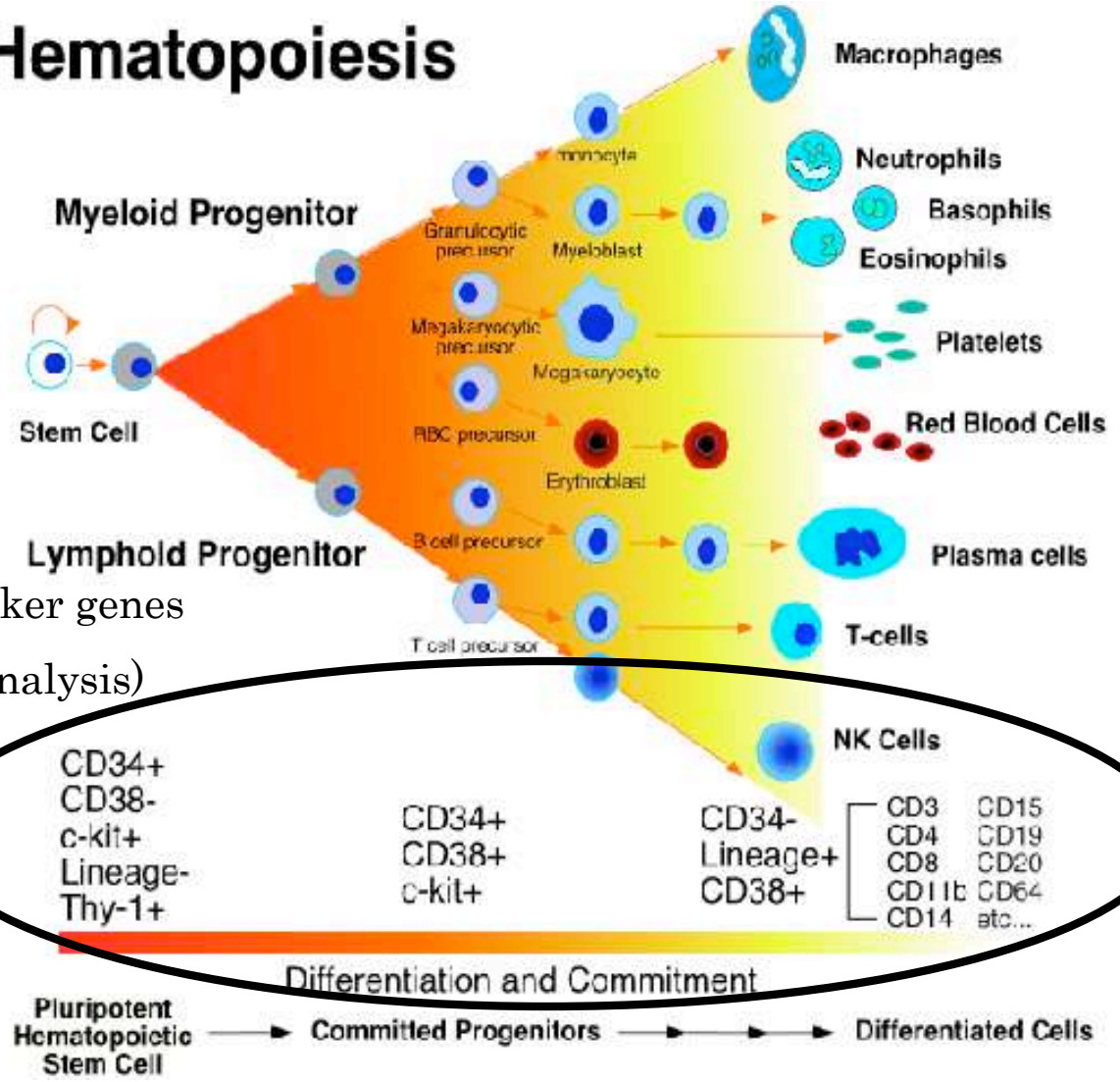
Classifier compendium
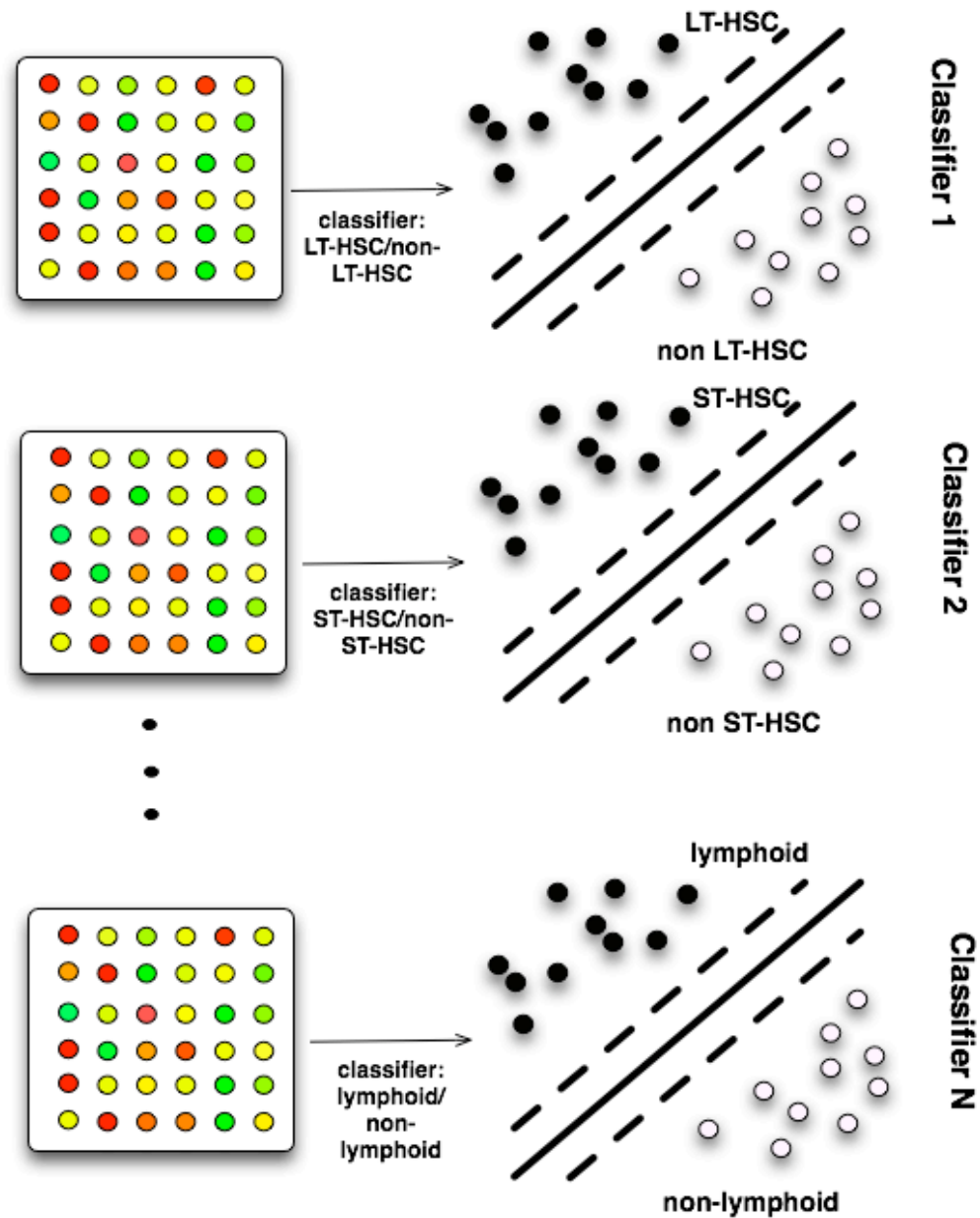
Summary and proposed steps for classification aim

- Complementarity of feature types
  - Feature selection
- Compendium of classifiers from stem cell differentiation experiments
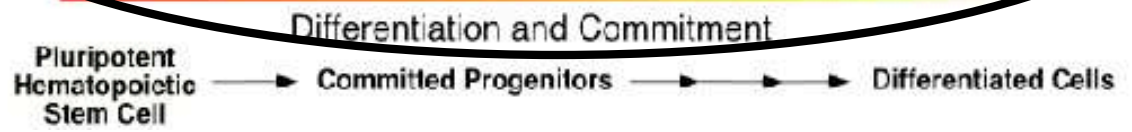- Evaluation
  - Hematopoietic system
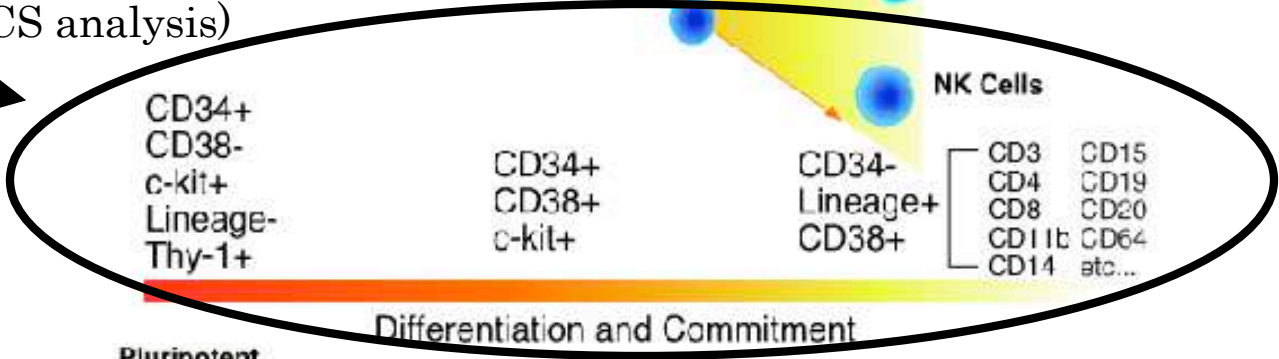
Current methods for stem cell population isolation and purification

Cell surface marker genes

(used in FACS analysis)
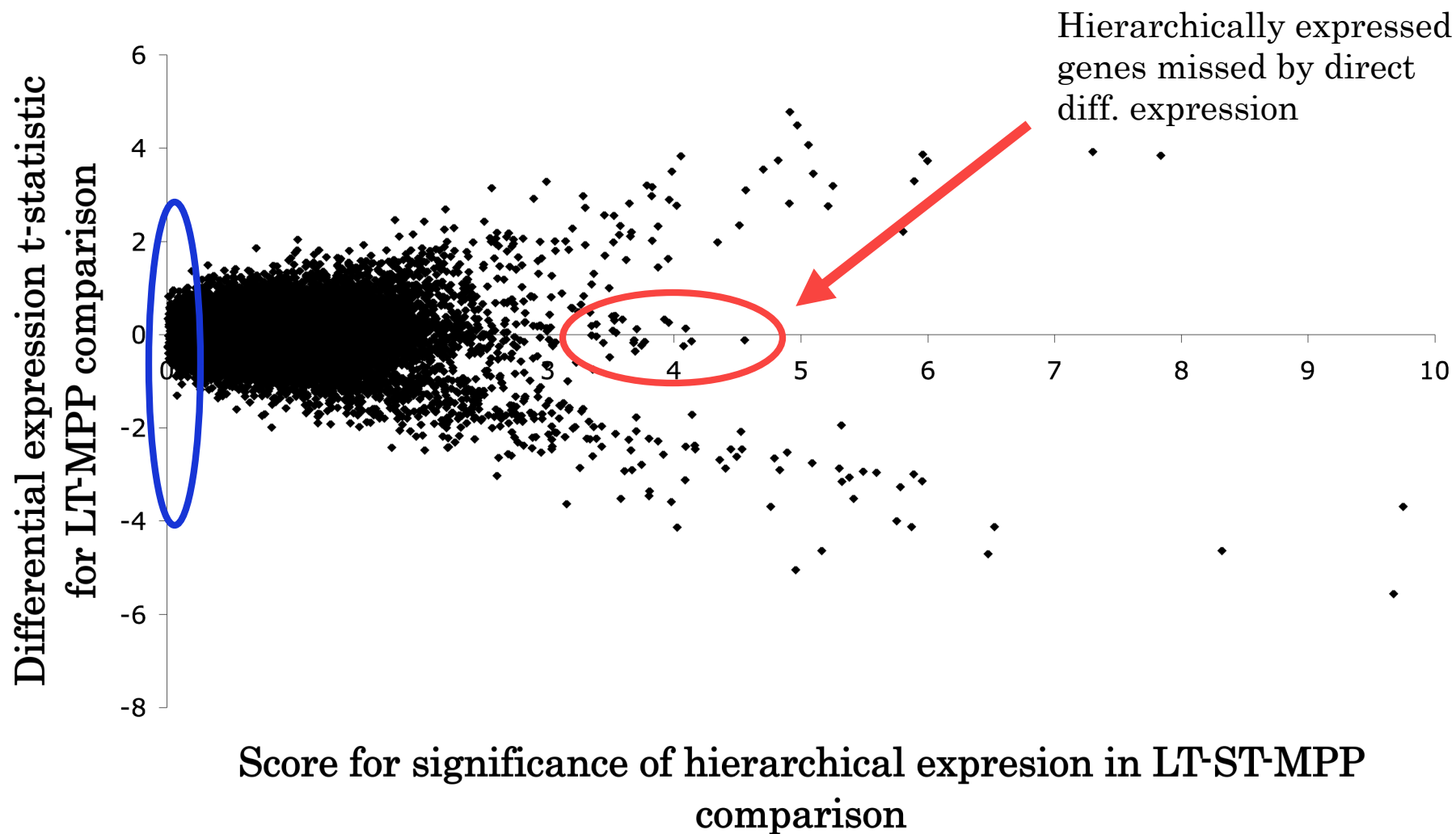
Aim 3: Systematic identification of hierarchically expressed genes

- Can we identify other indicator genes?
- Differential expression analysis
  - Hematopoietic system
  - ANOVA FDR-based approach

- Next step: hierarchical expression analysis

Scoring method for identifying indicator genes
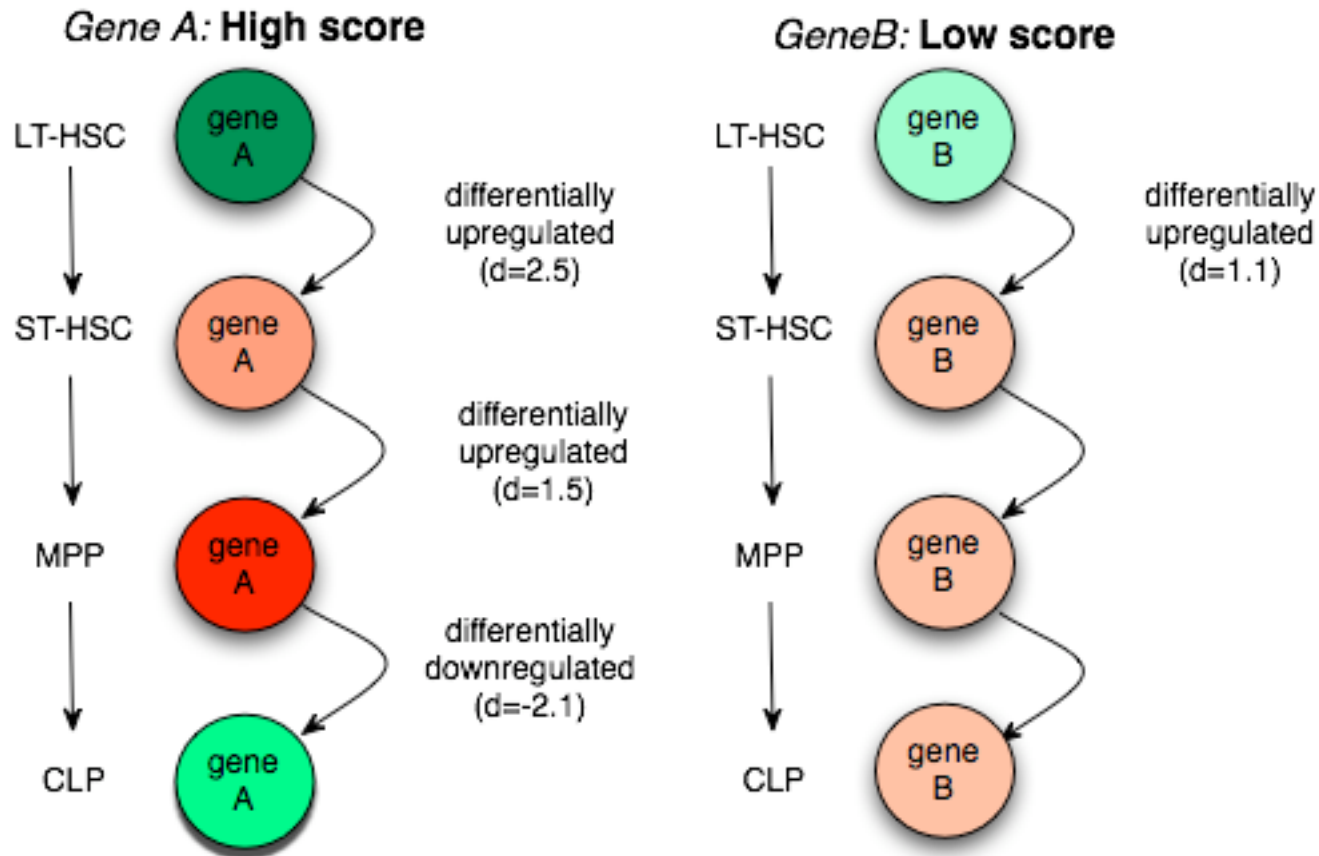
Can hierarchically expressed genes be missed by direct differential expression analysis?

# Summary and proposed steps for hierarchical expression detection analysis

- Method for identification of hierarchically expressed genes
- Apply to gene expression experiments with hierarchical stem cell lineages

Acknowledgements

- Josh Stuart
- Committee members
  - Kevin Karplus
  - Raquel Prado
  - Camilla Forsberg
- Collaborators
  - Weissman lab
  - Daniel Sam

- Others
  - Alex Williams
  - Charlie Vaske
  - Craig Lowe
  - David Bernick
  - Matt Weirauch

MicroRNA targets with inverse correlation: functional enrichment

Chromosome 17
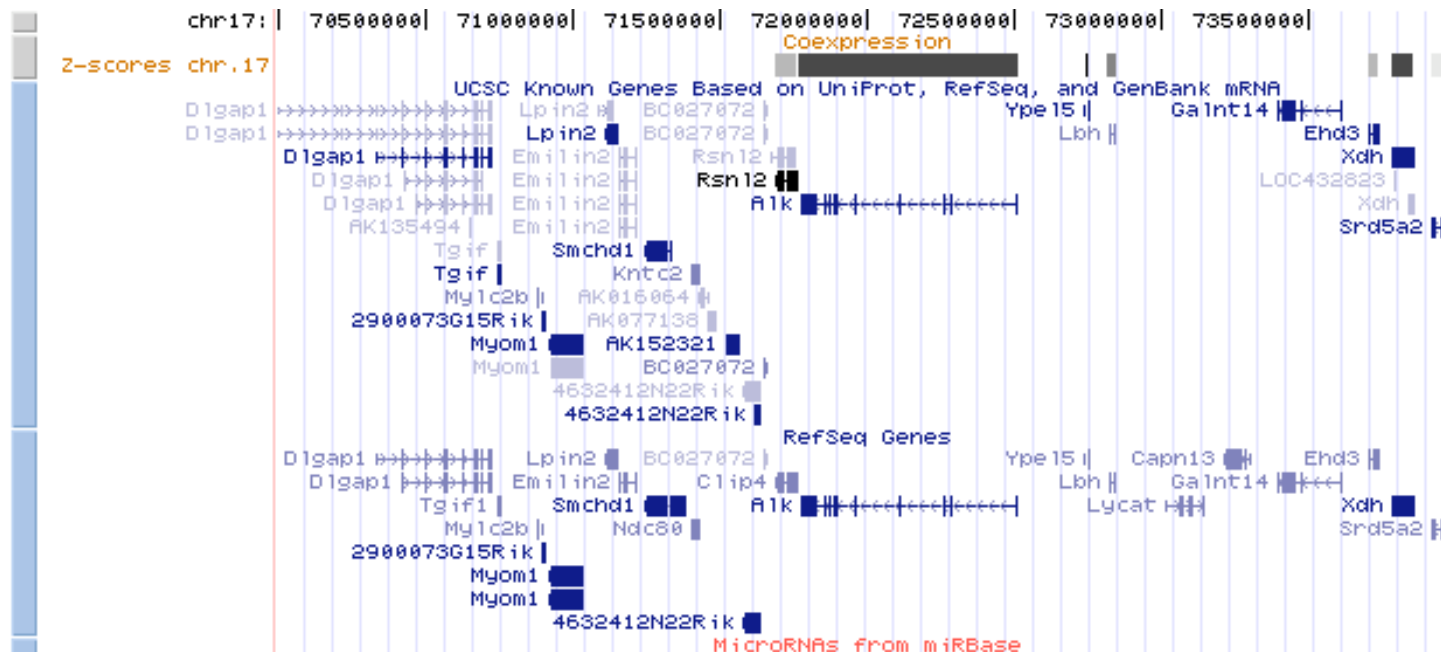
Alk - anaplastic lymphoma kinase: tyrosine kinase (orphan receptor; plays an important role in normal development
Xdh - xanthine dehydrogenase; regulation of epithilial cell differentiation

Chromosome 17

Marcksl1 - MARCKS-like 1 -- high level of co-expression with neighboring genes
Hdac1 - histone deacetylase 1

Cell surface marker genes